Improving multitasking assessment in healthy older adults using a prop-based version of the breakfast task

Citation for published version:
Kosowicz, M & MacPherson, SE 2016, 'Improving multitasking assessment in healthy older adults using a prop-based version of the breakfast task', *Applied Neuropsychology: Adult*, vol. 24, no. 3, pp. 252-263. https://doi.org/10.1080/23279095.2015.1136310

Digital Object Identifier (DOI):
10.1080/23279095.2015.1136310

Published In:
Applied Neuropsychology: Adult

Improving multitasking assessment in healthy older adults using a prop-based version of the Breakfast task

Maria Kosowicz & Sarah E. MacPherson

Centre for Cognitive Ageing and Cognitive Epidemiology, Human Cognitive Neuroscience, Department of Psychology, University of Edinburgh, UK

Address correspondence to:
Sarah E. MacPherson
Human Cognitive Neuroscience
Department of Psychology, PPLS
University of Edinburgh
7 George Square
Edinburgh, UK
EH8 9JZ
Tel: +44 (0)131 650 9862
Fax. +44 (0)131 651 3230
E-mail: sarah.macpherson@ed.ac.uk

Word count: 6941 words

Abstract word count: 151 words

ABSTRACT

Computerized cognitive assessment is becoming increasingly more common in clinical neuropsychological assessment and cognitive neuropsychological research. A number of computerized tasks now exist to assess multitasking abilities which are essential for everyday tasks such as cooking, shopping or driving, but little is known about whether these tasks are appropriate for assessing older adults' multitasking. The present study directly compared age effects on multitasking when assessed using a computerized and a prop-based version of Craik and Bialystok's (2006) Breakfast task. Twenty participants aged 18 to 24 years and 20 participants aged 60 to 79 years were assessed on both versions of the Breakfast task. While age-related decrements in multitasking performance were found using the computerized task, significant age differences were not found on the majority of measures when the prop-based version was administered. The results suggest that age-related deficits in multitasking will be less when more contextualized, non-computer based tasks are used.

Keywords: multitasking, aging, Breakfast task, assessment, ecological validity

Improving multitasking assessment in healthy older adults using a prop-based version of the Breakfast task

In neuropsychological assessment, measures which reflect the demands imposed in everyday life are thought to be better predictors of an individual's performance in daily life (Chaytor & Schmitter-Edgecombe, 2003). Therefore, it is important that assessment tools provide information about the level of functional difficulty that individuals may face (Spooner & Pachana, 2006). This is particularly important when assessing frontal executive abilities, given the dissociation reported in frontal patients who perform within normal limits on more traditional executive tasks administered in clinical assessments but perform poorly on everyday tasks in real-life (e.g., Shallice & Burgess, 1991).

Performance on everyday tasks such as preparing a meal (Craik & Bialystok, 2006), completing errands in a shopping center (Shallice & Burgess, 1991) or driving (Levy & Pashler, 2008) is typically assessed in clinical and research settings using tasks involving multitasking. Multitasking is the ability to complete several tasks within a limited time period by efficiently switching between the tasks and planning the best possible order to perform them. Individuals might not complete one task before starting another because of external time constraints, or because of physical or cognitive limitations (Burgess, 2000a, 2000b). There are obvious drawbacks in the administration of multitasking tasks in the real world such as being costly and time consuming to set up and the lack of experimental control over the external situation (e.g., Elkind, Rubin, Rosenthal, Skoff, & Prather, 2001). Therefore, multitasking has been assessed using prop-based board games, video recordings of real-world locations and computer programs such as virtual environments to simulate real-life settings (Rendell & Craik, 2000; Zhang et al., 2003; Craik & Bialystok, 2006; Farrimond, Knight, & Titov, 2006; Paraskevaides et al., 2010).

In the real world, older adults appear to be able to carry out everyday tasks such as preparing a meal or shopping for food with little difficulty (Phillips, Kliegel, & Martin, 2006). Indeed, prop-based tasks devised to assess multitasking such as the Six Elements task (Shallice & Burgess, 1991) or the Day Out task (Schmitter-Edgecombe, McAlister, & Weakley, 2012) have demonstrated that older adults perform as well as younger adults in terms of rule clarification and rule recall, and the number of tasks initiated and completed. However, the same studies show that older adults perform more poorly than younger adults when forming, retaining and executing plans, and rule breaking (Levine et al., 1998; Kliegel, McDaniel, & Einstein, 2000; Sanders & Schmitter-Edgecombe, 2012; McAlister & Schmitter-Edgecombe, 2013). These studies suggest that older adults can complete as many subtasks as the younger adults when multitasking but tend to do so less efficiently (Schmitter-Edgecombe et al., 2012).

Computerized assessment has been considered a suitable tool for assessing multitasking abilities and a number of computerized tasks now exist. For example, using a shopping mall virtual environment (VMALL), healthy older adults have been found to make significantly more overall and non-efficiency errors and perform more rule breaks than younger adults (Rand, Rukan, Weiss, & Katz, 2009). In another study, Craik and Bialystok (2006) administered a computerized cooking task which involved planning and prioritizing the cooking of various foods to ensure the foods were ready at the same time. They found age-related differences where older participants executed their plan less efficiently, often failing to perform appropriate actions at ideal times and making more perseverative errors than younger participants. More recently, the same research group have shown that the Breakfast task loads on an executive component and a non-executive cooking monitoring component and recommend the task as a quick, fun and realistic assessment of the cognitive abilities important for performing complex everyday tasks (Rose et al., 2015). Such laboratory-based computerized assessments of multitasking have reported age-related differences on the majority of measures examined.

The use of technology such as computers and virtual environments should be considered carefully by neuropsychologists when assessing multitasking. While older adults are now the fastest growing group of computer users in terms of internet access (Hart, Chaparro, & Halcomb, 2008; Medlock et al., 2012), more than half of adults aged ≥65 years do not use computers (Charness, Fox, & Mitchum, 2010). Computer acceptance is more difficult for older adults as they did not grow up with computers (Fozard & Wahl, 2012), and most will have never worked in a job where they used computers or were trained in computer skills (Jayroe & Wolfram, 2012). Older adults may find computer-based multitasking environments unfamiliar and artificial. It should also be considered that while computerized versions of tests may have outstanding face validity, their ecological validity may not be any better than standard paper and pencil neuropsychological instruments (Chan, Shumb, Toulopoulou, & Chend, 2008).

It is important to determine whether different versions of the same multitasking paradigm will result in similar findings. Previous studies that have adopted tasks mimicking real-life versus laboratory-based tasks, including our own work, (e.g., Garden, Phillips, & MacPherson, 2001; Schnitzspahn, Ihle, Henry, Rendell, & Kliegel, 2011; Sanders & Schmitter-Edgecombe, 2012; Hering, Cortez, Kliegel, & Altgassen, 2014) have tended to report age differences on both types of task. For example, Rand and colleagues (2009) examined multitasking in a real-life shopping mall environment and multitasking in a virtual shopping mall and found age differences on both versions of the task in terms of more partial, complete and non-efficiency errors in older adults. Only age effects in rule break errors were found in addition on the virtual but not the real-life shopping mall task. However, these paradigms are not necessarily matched in terms of structure, task demands or experimental control, making direct comparison between different versions difficult. One way to address this question is to compare computerized versus prop-based versions of the same test in the same individuals to determine whether similar age-related multitasking deficits are found using both methods of assessment.

The aim of the current study was to directly compare multitasking performance on a computerized and prop-based version of Craik and Bialystok's (2006) Breakfast task. A prop-based Breakfast task was devised which only differed in that it involved cardboard versions of the food, plastic cutlery, paper plates and kitchen timers. In the original computerized task, the complexity of the task varied depending on whether the subgoals were presented on the same (1-screen) or different (2- or 6-screen) computer screens. Analogous conditions were created in the prop-based version where the timers were either uncovered (1-display), or covered as a group (2-display) or individually (6-display). In line with Craik and Bialystok (2006), it was hypothesized that older adults would perform the computerized Breakfast task significantly more poorly than the younger adults. As the prop-based task removes the need to interface with virtual representations of cooking-related objects, it is predicted older adults will show improved performance on this version.

METHOD

Participants

Twenty adults (19 right-handed, 11 women) aged between 18 and 24 years (M = 21.25, SD = 1.74) and 20 right-handed adults (13 women) aged between 60 and 79 years (M = 67.00, SD = 5.68) were tested. These particular age ranges were included in an attempt to replicate the original age effects reported by Craik and Bialystok (2006). The two age groups did not differ significantly in their number of years of full-time education (M = 15.15, SD = 1.46, range = 12-17; M = 15.50, SD = 2.84, range = 12-22 for younger and older adults respectively), $t(38) = 0.49$, $p = .63$. Participants were volunteers recruited through the university, word-of-mouth via other volunteers, or were known to the experimenter. Inclusion criteria included: aged between 18 and 30 years or 60 and 80 years, normal or corrected to normal hearing and vision, living independently in the community, a score of 82 or above (out of 100) on the Addenbrooke's Cognitive Examination-Revised (ACE-R; Mioshi, Dawson, Mitchell, Arnold, & Hodges, 2006) and no self-reported history of the neurological or psychiatric disorders listed on the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III UK) and the Wechsler Memory Scale-Third Edition (WMS-III UK) selection criteria (Wechsler, 1997a, 1997b). All participants gave written, informed consent in accordance with the Department of Psychology Research Ethics Committee guidelines.

Measures

Background measures. Participants were administered the following paper-and-pencil measures: Addenbrooke's Cognitive Examination-Revised (ACE-R) to assess overall cognitive abilities (Mioshi et al., 2006); digit symbol substitution test to assess speed of processing (Wechsler, 1981); digit span forwards to assess short-term verbal memory, and digit span backwards (Wechsler, 1981) and the alpha span test (Craik, 1986) to assess working memory. Similar measures of processing speed and working memory (as well as fluid intelligence and inhibition) were found to load on a general executive factor by Rose et al. (2015) with better executive performance associated with better performance on certain Breakfast task measures.

Computerized Breakfast task. The computerized Breakfast task was devised by Craik and Bialystok (2006). The task was presented on a laptop computer with a screen of 34cm x 21cm. An ergonomic input device was used in all cases to facilitate movement and increase comfort given that older adults may be less familiar with a standard computer mouse. This was in contrast to Craik and Bialystok (2006) who used a touch-screen monitor. Participants were presented with five foods on the computer screen and the times necessary to cook them. The task was to prepare a virtual breakfast so the foods were ready at the same time, as well as setting a table while the foods were cooking. The cooking times were different for each food, ranging from 2 to 5.5 minutes. The “start cooking” button was placed above each food name and the stop cooking button was below the name. Once the “start cooking” button was pressed, the name of the food was highlighted in green and after the stop button was pressed, it changed to red. There was also a timer visible under each food icon which indicated the state of cooking. The timer was activated after pressing the food’s start button and appeared as a column of time decreasing in height every second while the food was cooking, from its start time (e.g., 5.5 minutes) to zero. There was no visual or auditory indication of the overcooking time. Therefore, participants had to monitor the cooking time themselves and remember to switch off the foods once the timers reached zero.

While the foods were cooking, participants were required to set a table for four people. This was a distracter task aimed to increase the planning and prospective memory demands of the task. Cutlery (forks, knives, spoons) and plates were to be placed in the correct positions on the table. Participants had to select an item by clicking on it and then click on its new position. There were no strict rules for the table setting: participants could complete a whole place setting and then continue to the next one or set all plates or particular items of cutlery first and then continue to set other items in the place setting. It was, however, possible to move only one item at a time. Once all four table settings were complete, the items returned to their initial positions and the table setting task started again. The number of successful table settings was recorded. Participants were informed that the main goal of the task was to cook the foods so that they were ready as close together as possible while table setting was the secondary goal of the task.

The Breakfast task involved three conditions with different levels of complexity. In the 1-screen condition, the timers for all five foods and table settings were visible on the same screen. The more complex 2-screen condition required participants to keep in mind the progress of the cooking and switch between a screen displaying the timers for all foods and a second screen displaying the table to be set. While participants were setting the table, the only indicator of the cooking status was information that the cooking of a particular food had not been started yet (no highlight of the food name), it had been started (green highlight) or it had been already stopped (red highlight). By simply pressing one of the food icons, the participants could switch from the table setting screen to the food screen in order to start or stop food cooking, or check their progress. To return to setting the table, participants had to press the return to table button. The 6-screen condition involved five screens displaying the timers specific to each food and one screen displayed the table to be set. As in the 2-screen condition, the screen with the table settings displayed basic information about the foods but did not show the passing time. Participants could switch between the screens in the same way by pressing a food icon. However, to switch to another food, they had to return to the table setting screen and then press the button for that food.

The actions performed by participants were automatically recorded by the computer program. Direct measures included the starting and stopping times for the cooking of the foods, the range between the stop times for the first and the last food cooked, the number of table places set and the frequency of cooking progress checks in the 2-screen and 6-screen conditions. Indirect measures included the discrepancy between the real and ideal start and stop times and the amount of time spent on setting the table when the foods should have been started or stopped relative to the ideal times. A practice task involving two foods was performed prior to the actual Breakfast task.

Prop-based Breakfast task. The prop-based Breakfast task used the same five foods made out of cardboard (see Figure 1). As in the computerized version (see Figure 1 of Craik & Bialystok, 2006), the names of the foods and required cooking times were displayed. Instead of timers in the form of columns, there were digital kitchen timers next to the names of each food. Participants were asked to cook the foods so they were ready around the same time and set as many table places as they could while the foods were cooking. Participants started cooking by pressing the start buttons on the timers and stopped cooking by pressing the stop buttons on the same timers. Another digit kitchen timer was available to the experimenter to note down when the cooking of each food was started and stopped.

- Insert Figure 1 around here -

To avoid the timers bleeping once they reached zero, they were set for 10 minutes longer than the actual cooking time. Participants were not aware of this, since the first digit on the display (i.e. “1”) was covered with a small piece of paper displaying the food icon. This also allowed the experimenter to record the overcooking time.

For the table setting task, disposable plates and cutlery were used. Seventy plastic forks, 70 knives, 70 spoons and 70 paper plates were laid out on the table. Participants were asked to place the plates in a pile in front of them and put the cutlery in separate boxes provided. Participants were asked to only move one plate or one piece of cutlery at a time.

As in the computerized task, three analogous conditions differing in complexity were performed. In the 1-display condition, the food cooking and setting of the table were conducted at the same table with all food timers uncovered. In the 2-display condition, the timers for the foods were covered with a long box while participants set the table. Whenever the participant wanted to start or stop cooking or check the foods, they had to lift the long box to make all the timers visible. The experimenter noted down every time a box was lifted to check the time. To go back to setting the table, participants had to cover the timers again. In the 6-display condition, each timer was covered with a separate box so that only one timer could be seen at a time. All timers had to be covered while participants were setting the table. When participants wanted to start, stop or check the progress of cooking, they could uncover the boxes one at a time. They had to put the box back if they wanted to uncover the second timer or go back to setting the table.

The start and stop cooking times for each food and progress checks were recorded manually by the experimenter. The number of complete places set was also counted at the end of the task. The direct and indirect measures recorded were identical to those in the computerized version of the task. A 2-item practice task was performed first.

Procedure

Firstly, the participants completed the background tests (approximately 40 minutes). The study had a repeated measures design where all participants performed both experimental conditions. Participants were randomly allocated to one of two groups to determine in which order they performed the computerized and the prop-based Breakfast tasks. Half the younger and older participants performed the computerized version first and the other half performed the prop-based version first (approximately 20 minutes for each version). However, the 1-screen/display condition was always performed first followed by the 2- and 6-screen/display conditions as Craik and

Bialystok (2006) noted that older participants had difficulties understanding the task if the 6-screen condition was performed first. Participants were allowed to take breaks as and when they needed them.

Data analyses

The performance of the younger and older adults on the background measures and table setting times was compared using independent samples t-tests. The one-screen/display and two-screen/display conditions were considered practice trials for the six-screen/display conditions (Rose et al., 2015). Data for the 6-screen/display conditions are presented as they require the most monitoring, planning and task-switching. A 2 (age) x 2 (task) analysis of variance (ANOVA) was conducted for each of the Breakfast task measures; age was a between-subjects variable whereas task was a within-subjects variable. If the data were not normally distributed, they were log transformed. The Shapiro-Wilk test was conducted to test the assumption that the data were normally-distributed with the p-values reported in Tables 2 and 3. Post-hoc independent samples t-tests and paired samples t-tests were Bonferroni adjusted for multiple comparisons although uncorrected significance levels are reported. G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) was used to compute the effect size for the two-way interactions in the ANOVAs using a sample size of 40 participants with 20 participants in each age group and 90% power at the 5% alpha level. The analysis revealed a medium effect size of 0.26 (Cohen, 1988; Faul et al., 2007).

RESULTS

Background measures

The performance of the younger and older adults on the background measures is shown in Table 1. Older adults had significantly poorer scores than younger adults on the ACE-R, $t(38) = 4.50$, $p < .01$, digit span forward, $t(38) = 4.37$, $p < .01$, digit span backward, $t(38) = 3.53$, $p < .01$, digit symbol substitution, $t(38) = 7.27$, $p < .01$, and alpha span, $t(38) = 3.92$, $p < .01$. For the ACE-R test, none of the participants performed below the cutoff of 82.

- Insert Table 1 around here -

Cooking and table setting measures

Discrepancy. The discrepancy measure assesses the ability to stop cooking the food at the correct time and therefore denotes under- or over-cooking. It was calculated using the following formula: *discrepancy = |actual cooking time – ideal cooking time|*. Since the absolute value of the difference was taken, all discrepancy scores were above zero, indicating when foods were not switched off on time but not whether it was too early or too late. The mean discrepancy scores and standard deviations for the two groups performing the computerized and prop-based Breakfast tasks are shown in Tables 2 and 3.

As the discrepancy data were not normally distributed, they were log transformed. A 2 (age) x 2 (task) ANOVA revealed the discrepancy score was significantly higher in the older group, $F(1, 38) = 25.67$, MSE = 1.88, $p < .0001$, $\eta_p^2 = .40$. The task main effect showed the prop-based task was performed better than the computerized one, $F(1, 38) = 11.27$, MSE = 0.62, $p < .005$, $\eta_p^2 = .23$.

There was also a significant age x task condition, $F(1, 38) = 13.56$, MSE = 0.74, $p < .005$, $\eta_p^2 = .26$, where the older adults had significantly higher discrepancy scores when the task was computerized ($p < .0001$) compared to younger adults but not when it was prop-based ($p = .18$). The younger participants did not differ in their discrepancy scores across the two versions ($p = .72$) whereas the older adults had better discrepancy scores when the task was prop-based ($p < .005$).

Range of stop times. The range of stop times denoted the difference in the times between the stopping of the last and first foods. It was calculated: *range = time last food stopped – time first food stopped.* Ideally this measure should approach zero. The mean range of stop times and standard deviations for each age group are shown in Tables 2 and 3. The range of stop times were not normally distributed and were therefore log transformed.

The 2 x 2 ANOVA revealed the range was significantly higher in the older group, $F(1, 38) = 24.50$, MSE = 1.44, $p < .0001$, $\eta_p^2 = .39$. There was also a significant main effect of task, $F(1, 38) = 18.89$, MSE = 1.20, $p < .0001$, $\eta_p^2 = .33$, where the prop-based task had a significantly lower range score compared to the computerized version.

The age and task interaction approached significance, $F(1, 38) = 3.91$, MSE = 0.25, $p = .06$ $\eta_p^2 = .09$, where older adults had significantly higher range scores than the younger adults on the computerized task ($p < .001$) but the two groups did not significantly differ in terms of their range scores on the prop-based version ($p = .05$).

- Insert Tables 2 and 3 around here -

Deviation from ideal start times. Table 4 shows the formulae for calculating the ideal start times for each food. The actual start time was then subtracted from the ideal start time for each food to provide a deviation from the ideal start time. The average deviation between the ideal and actual start times for Foods 2-5 was then calculated. The mean deviation from ideal start times and standard deviations for each age group are demonstrated in Tables 2 and 3. This measure indicated how precisely participants were able to start the foods so that they finished cooking as closely together as possible. Preferably, there should be no deviation from the ideal start times. However, if the participant forgot to start a food at its ideal time, the ideal start time for the next food changed to a later time so that the food could be ready relatively close to the first food's start time but also relatively close to the food that had been started too late.

Firstly, the data were log transformed as they were not normally distributed. A 2 ~~x 3~~ x 2 ANOVA demonstrated the deviation score was significantly higher in the older than the younger group, $F(1, 38) = 13.37$, MSE = 0.87, $p < .005$, $\eta_p^2 = .26$. The main effect of task, $F(1, 38) = 26.17$, MSE = 1.31, $p < .0001$, $\eta_p^2 = .41$, demonstrated the participants had a significantly higher deviation score in the computerized version of the task.

The age x task interaction approached significance, $F(1, 38) = 3.92$, MSE = 0.20, $p = 0.06$, $\eta_p^2 = .09$, where the younger adults had a significantly smaller deviation score than the older adults in the computerized version of the task ($p < .0001$) but the two groups did not differ in the prop-based version ($p = .19$). The younger group did not significantly differ in terms of their deviation score across the two versions of the task ($p = .03$) whereas the older age group had a significantly lower deviation score in the prop-based task compared to the computerized one ($p < .0001$).

Checking the progress of cooking. The mean number of times that participants checked the cooking progress and the standard deviations are shown in Tables 2 and 3. A 2 (age) x 2 (task) ANOVA revealed that the main effects of age, $F(1, 38) = 3.85$, MSE = 20.00, $p = .06$, $\eta_p^2 = .09$, and task, $F(1, 38) = 2.16$, MSE = 13.89, $p = .15$, $\eta_p^2 = .05$, were not significant.

There was a significant age x task interaction, $F(1, 38) = 6.09$, MSE = 39.20, $p < .05$, $\eta_p^2 = .14$, where younger adults checked the cooking progress significantly more than the older adults in the computer version ($p < .01$) but not the non-computer version of the task ($p = .58$).

Number of table settings completed. The mean number of table settings and standard deviations for the two age groups are shown in Tables 2 and 3. The age x task ANOVA conducted for the number of table settings completed demonstrated the younger group completed significantly more table settings than the older group, $F(1, 38) = 98.70$, MSE = 10757.07, $p < .0001$, $\eta_p^2 = .72$. The main effect of task was also significant, $F(1, 38) = 90.17$, MSE = 7175.74, $p < .0001$, $\eta_p^2 = .70$, where more table settings were completed in the prop-based version of the task.

The two-way interaction between age and task was significant, $F(1, 38) = 8.14$, MSE = 647.90, $p < .01$, $\eta_p^2 = .18$, where younger adults completed significantly more table settings than older adults in both the computerized and prop-based tasks ($p < .0001$).

Controlling for age-related cognitive decline

The same 2 x 2 ANOVAs were conducted for each of the Breakfast task measures but controlling for performance on the background tests. The analysis of covariance revealed that the two-way age x task interaction remained significant for checking the progress of cooking, $F(1, 33) = 4.68$, MSE = 32.19, $p < .05$, $\eta_p^2 = .12$ and the discrepancy score approached significance, $F(1, 33) = 4.22$, MSE = 0.25, $p = .05$, $\eta_p^2 = .11$. However, the two-way interactions for the deviation from ideal start times, $F(1, 33) = 0.58$, MSE = 0.03, $p = .45$, $\eta_p^2 = .02$, range of stop times, $F(1, 33) = 2.35$, MSE = 0.16, $p = .14$, $\eta_p^2 = .07$, and number of table settings completed, $F(1, 33) = 1.26$, MSE = 84.67, $p = .27$, $\eta_p^2 = .04$, were no longer significant.

Time spent performing Breakfast task

Table 5 demonstrates the amount of time spent performing the computerized and prop-based versions of the Breakfast task. The data were log transformed when they were not normally distributed and 2 (age) x 2 (task) ANOVAs conducted. Older adults spent more time than younger adults on the computerized version ($p < .01$) of the Breakfast task but a similar a length of time on the prop-based version ($p = .02$). Both groups spent more time on the computerized version than the prop-based version (younger: $p < .005$; older: $p < .0001$), $F(1, 38) = 4.72$, MSE = 0.10, $p < .05$, $\eta_p^2 = .11$. There was also a significant main effect of age, $F(1, 38) = 10.65$, MSE = 2771.39, $p < .005$, $\eta_p^2 = .22$, and task, $F(1, 38) = 14.99$, MSE = 0.10, $p < .0001$, $\eta_p^2 = .28$.

In terms of the length of time spent on table setting, both age groups spent more time table setting during the prop-based task, $F(1, 38) = 5.85$, MSE = 0.06, $p < .05$, $\eta_p^2 = .13$. The main effect of age, $F(1, 38) = 0.002$, MSE = 0.10, $p = .64$, $\eta_p^2 = .01$, and the two-way interaction were not significant, $F(1, 38) = 0.03$, MSE = 0.000, $p = .86$, $\eta_p^2 = .001$. For time spent cooking, older adults spent more time on the cooking task than younger adults when performing the computerized version ($p < .001$) but a similar a length of time on the prop-based version ($p = .02$). Both groups spent more time on the computerized version than the prop-based version (younger: $p < .01$; older: $p < .0001$), $F(1, 38) = 15.29$, MSE = 0.93, $p < .0001$, $\eta_p^2 = .29$. There was also a significant main effect of age, $F(1, 38) = 200.76$, MSE = 12.25, $p < .0001$, $\eta_p^2 = .84$, and task, $F(1, 38) = 31.62$, MSE = 3.17, $p < .0001$, $\eta_p^2 = .45$.

Older adults spent a significantly smaller proportion of the total time setting the table compared to younger adults on the computerized version ($p < .0001$) of the Breakfast task but a similar a length of time on the prop-based version ($p = .06$). Both age groups spent a significantly smaller proportion of the time setting the table on the computerized version compared to the prop-based task (both $p < .0001$), $F(1, 38) = 17.65$, MSE = 0.05, $p < .0001$, $\eta_p^2 = .32$. There were also main effects of age, $F(1, 38) = 25.14$, MSE = 0.09, $p < .0001$, $\eta_p^2 = .40$, and task, $F(1, 38) = 103.47$, MSE = 0.31, $p < .0001$, $\eta_p^2 = .73$. Older adults, compared to younger adults, spent a greater proportion of time setting the table when they should have been starting or stopping the food both on the computerized ($p < .0001$) and the prop-based ($p < .005$) tasks. However, older adults spent significantly less time inappropriately table setting on the prop-based task compared to the computerized task ($p < 0.05$). The younger adults did not differ between tasks ($p = .78$), $F(1, 38) = 7.88$, MSE = 0.09, $p < .01$, $\eta_p^2 = .17$. Both main effects were also significant (age: $F(1, 38) = 31.19$, MSE =. 37, $p < .0001$, $\eta_p^2 = .45$, and task: $F(1, 38) = 6.99$, MSE = 0.08, $p < .05$, $\eta_p^2 = .16$).

- Insert Table 5 around here -

DISCUSSION

Our results demonstrated that while older adults performed significantly more poorly than younger adults on the computerized version of the Breakfast task, significant age-related differences were not found in the same participants on the majority of measures on the analogous prop-based task. Not only were the older adults less likely to under- or over-cook the food when the computer element of the task was removed, they were also more able to start cooking the different foods so they would be ready around the same time and stop the foods cooking at the same time. To our knowledge, this is the first study to directly compare the ability of the same group of older adults to multitask using computerized and prop-based versions of the same paradigm. While our prop-based Breakfast task is clearly not a real-life task given that the foods are cardboard and no real cooking is required, it does show that removal of the computer component improves older adults' performance. Previous work, including our own, has not reported age-related differences on certain aspects of multitasking when more naturalistic versions of tasks are used (Garden et al., 2001; Levine et al., 1998). Further work might explore whether performance on the prop-based versions of the Breakfast task or other multitasking paradigms might predict performance in real-life situations.

While the current study clearly shows this marked improvement in multitasking performance in older adults when a prop-based version of the Breakfast task is administered, it cannot answer the question of why the older adults perform this version so much better. It may be that older adults are more apprehensive about using computers or older adults are not familiar with using them. Younger and older adults' experience and ability to use computers clearly still differs (e.g., Pew Internet and American Life Project, 2004; Dickinson, Eisma, Gregor, Syme, & Milne, 2005; Charness et al., 2010) and studies have shown that older adults have lower beliefs in their computer abilities (Hawthorn, 2007; Mitzner et al., 2010) and more computer anxiety than younger adults (Czaja et al., 2006; Turner, Turner, & Van De Walle, 2007; Wagner, Hassanein, & Head, 2010). In turn, this may mean that older adults are less motivated to perform well on the computerized Breakfast task due to their low self-efficacy. Regrettably our current study did not record attitudes towards computers or familiarity. However, previous research comparing the performance of older adults on computerized and paper and pencil versions of tasks assessing memory, attention and speed of processing have actually demonstrated that older adults were less likely to rate the computerized versions of the tests as "difficult, stressful, or unacceptable" compared to pencil-and-paper ones (e.g., Collerton et al., 2007, p. 1634). Furthermore, a recent large-scale study comparing older adults performing computerized and paper and pencil tests of processing speed, attention and executive function found that individuals with better computer experience tended to perform better on cognitive tests regardless of whether they were computerized or not (Fazeli, Ross, Vance, & Ball, 2013). These findings suggest that computer experience alone is not sufficient to explain older adults' performance on computerized cognitive measures.

There is also the issue of familiarity with the input device used. Craik and Bialystok (2006) used a touchscreen. As a touchscreen was not available at the time of testing, an ergonomic mouse was used to facilitate movement and increase comfort in older adults. In the area of human-computer interaction, Schneider, Wilkes, Grandt and Schlick (2008) compared various input devices in 90 individuals aged 20 to 72 years and found that older adults were significantly poorer at performing tasks using a touchscreen compared to using a mouse. However, other aging studies involving mouse use compared to other input devices such as trackballs and light pens have found that older adults perform more poorly using a mouse (e.g., Chaparro, Bohan, Fernandez, Choi, & Kattel, 1999; Charness, Holley, Feddon, & Jastrzembski, 2004; Ketcham, Seidler, Van Gemmert, & Stelmach, 2002). Further work is required to understand the relationship between computer technology and multitasking performance.

Differences between laboratory-based experiments and real-world situations may exist for reasons in addition to the computerized nature of laboratory-based tasks. Often real-life tasks are habitual and so individuals can draw upon their past experience associated with that particular context. Real-life problems may still exist in novel situations (Craik, personal communication). Furthermore, individuals may be more motivated to perform real-life activities such as cooking a meal for friends, as they see these tasks as having more bearing on their lives compared to laboratory tasks. Real-world tasks also allow individuals to use external aids such as diaries, notepads and timers to help them successfully achieve their overall goal. Finally, within the laboratory environment, participants are intensely aware that their performance is being monitored, recorded and evaluated. Research within psychology has stressed the effects of being a participant within a study, the influence of an experimental investigator as well as the situation on performance (e.g., Rosenthal, 1967).

Of course there are alternative methods in the literature that attempt to ensure age-related biases do not contaminate performance on computerized assessments. For example, staircase algorithms are commonly used where the stimulus values presented essentially depend on an individual's prior responses (for a review see Leek, 2001). Performance on the task starts at a comparatively easy level but increases in difficulty if individuals perform well and declines if individuals perform poorly. This continues until the level of performance stabilizes and the limit of each person's performance can be determined. Indeed, in our own work investigating dual task abilities in healthy aging, we have consistently shown that younger and older adults do not significantly differ in their dual task performance when both tasks are performed at an individual's own ability levels (e.g., MacPherson, Della Sala, & Logie, 2004; MacPherson et al., 2007). In contrast, studies that do not equate performance across age groups tend to report age effects (Craik, 1977; McDowd & Craik, 1988; Salthouse, Rogan, & Prill, 1984; Wright, 1981).

It should also be noted that the lack of significant differences in performance between the age groups on the prop-based task does not necessarily indicate that older adults are not impaired on the cognitive abilities underlying multitasking. Indeed, in the present study, our older adults did show impairments on the background tests measuring speed of processing, short-term memory and working memory. When performance on these background measures was controlled for, a significant age x task interaction only remained for the checking the progress of cooking. Therefore, cognitive abilities do underlie performance on some of the computerized Breakfast task measures.

One explanation why some of the two-way interactions remained even when cognition was controlled for might be older adults' better knowledge of cooking. In the prop-based task, older adults may benefit from being more experienced and better cooks and so they check the foods less. While this would strengthen the claim that the prop-based task better reflects multitasking abilities in real-life (at least in terms of cooking), unfortunately this information was not recorded. Another limitation of the study is not having data to demonstrate that the younger and older adults' were matched in terms of fluid intelligence. Needless to say, years of full-time education might be considered as a crude measure of fluid abilities and the two age groups did not differ. Overall, the results of the paper and pencil prop-based task suggest-that, despite these declines in cognitive abilities, older adults can perform in multitasking environments, at least in terms of the Breakfast task, if the paradigm does not require older adults to interact with computers.

In terms of the table setting task, older adults completed significantly fewer table settings regardless of whether the task administered was computerized or prop-based. Older adults may have neglected this task given the emphasis put on cooking the foods to be ready on time. Another explanation might be reduced processing speed or the motor component involved in physically sorting the plates and cutlery. Declines in motor control have been reported in healthy adult aging (Smith, Sharit, & Czaja, 1999) which include coordination difficulties (Seidler, Alberts, & Stelmach, 2002), more erratic movements (Darling, Cooke, & Brown, 1998) and slowing of movement (Diggles-Buckles, 1993). These motor deficits have a negative impact on older adults' ability to perform activities in everyday life, which might include tasks such as setting the table. It is unlikely, however, that deficits in motor control can solely explain the significantly fewer table settings achieved by older adults. Rose et al. (2015) found that table settings significantly correlated with performance on measures that loaded on an executive factor including measures of working memory and processing speed. In the current study, table settings in both Breakfast tasks also significantly correlated with measures of working memory and processing speed.

Older adults spent more time overall, as well as cooking (at the expense of table setting), on the computerized Breakfast task, and yet performed the task more poorly than younger adults. The only measure not to show improved older adult performance on the prop-based Breakfast task was inappropriate table settings. Older adults continued to set the table when the food should have been started or stopped cooking regardless of the Breakfast task administered. The failure to stop setting the table when older adults should have been cooking has been attributed to perseverative behavior (Craik & Bialystok, 2006), which in turn might be associated with age-related mild frontal lobe decline (Grady & Craik, 2000; West, 1996).

A further limitation of our study relates to the scoring of the Breakfast task. Difference scores have been criticized for having low reliability compared to the reliability of the scores on which the difference score is based. However, difference scores continue to be widely used in behavioral research, as it has been argued that low difference score reliability does not necessarily mean low statistical power for mean comparisons (Haertel, 2006).

Other researchers have developed virtual environments to study multitasking and similar performance has been reported when real-life and virtual versions of the same task have been compared (McGeorge et al., 2001). Therefore, virtual environments may be an appropriate, safer, and better controlled setting for assessment of multitasking abilities. Computerized versions of virtual tasks have been developed to assess multitasking such as the Edinburgh Virtual Errands Test, Virtual Mall and Virtual Street (e.g., Logie, Trawley, & Law, 2011; McGeorge et al., 2001; Rajendran et al., 2011; Rand, Rukan, Weiss, & Katz, 2009). However, while these virtual environments have been used to assess multitasking in neurological patients (e.g., Jovanovski et al., 2012; McGeorge et al., 2001; Rajendran et al., 2011; Titov & Knight, 2005), little is known about healthy older adults' performance using them. Rand et al. (2009) demonstrated that older adults broke more rules and were less efficient than younger adults when a virtual supermarket environment was used. However, again, this age effect might be explained in terms of the computer aspect of the task. Like computerized tasks, older adults tend not to be familiar with virtual environments and therefore older adults must receive training to navigate and interact with the environment prior to assessment. While multitasking in virtual environments might strike a good balance between experimental control and ecological validity (Law, Trawley, Brown, Stephens, & Logie, 2012), future work is required to compare the ability of older adults to use virtual environments in order to carry out multitasking and compare their performance with more real-life multitasking paradigms.

In summary, the current study indicates that the chosen paradigm may be crucial when assessing multitasking in older adults. Laboratory-based computerized methods bring possibilities for improved task design, as well as higher efficiency and speed of data recording. This, however, may be at the expense of computer bias across distinct groups who differ in their computer experience. When the multitasking paradigm does not involve the interaction with computers, older adults demonstrated improved multitasking abilities.

REFERENCES

Burgess, P. W. (2000a). Real-world multitasking from a cognitive neuroscience perspective. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance VVIII* (pp. 465-472). Cambridge, MA: MIT Press.

Burgess, P. W. (2000b). Strategy application disorder: the role of the frontal lobes in human multitasking. *Psychology Research*, *63*(3-4), 279-288.

Chan, R. C., Shum, D., Toulopoulou, T., & Chen, E. Y. (2008). Assessment of executive functions: review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology, 23(2),* 201-216.

Chaparro, A., Bohana, M., Fernandez, J., Choib, S., & Kattelb, B. (1999). The impact of age on computer input device use:: Psychophysical and physiological measures. *International Journal of Industrial Ergonomics*, *24*, 503-513.

Charness, N., Fox, M. C., & Mitchum, A. L. (2010). Lifespan cognition and information technology. In K. Fingerman, C. Berg, T. Antonnuci, & J. Smith (Eds.), Handbook of Lifespan Psychology. New York: Springer.

Charness, N., Holley, P., Feddon, J., & Jastrzembski, T. (2004). Light pen use and practice minimize age and hand performance differences in pointing tasks. *Human Factors, 46(3),* 373-384.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review,* 13, 181–197.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Collerton, J., Collerton, D., Arai, Y., Barrass, K., Eccles, M., Jagger, C., McKeith, I., Saxby, B. K., Kirkwood, T.; Newcastle 85+ Study Core Team. (2007). A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ Pilot Study. *Journal of the American Geriatrics Society, 55(10),* 1630-1635.

Craik, F. I. (1977). Age differences in human memory. In J. E. Birren & K.W. Schaie KW (Eds.), *The Handbook of the Psychology of Aging* (pp. 384-420). New York: Van Nostrand Reinhold.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human Memory and Cognitive Capabilities: Mechanisms and Performances* (pp. 409-422). Amsterdam: Elsevier.

Craik, F. I. M., & Bialystok, E. (2006). Planning and task management in older adults: cooking breakfast. *Memory and Cognition, 34*(6), 1236-1249.

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A., & Sharit J. (2006). Factors predicting the use of technology: Findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and Aging, 21*(2), 333-352.

Darling, W. G., Cooke, J. D., & Brown, S. H. (1998). Control of simple arm movements in elderly humans. *Neurobiology of Aging, 10*(2), 149-157.

Dickinson, A., Eisma, R., Gregor, P., Syme, A., & Milne, S. (2005). Strategies for teaching older people to use the World Wide Web. *Universal Access in the Information Society, 4,* 3–15.

Diggles-Buckles, V. (1993). Age-related slowing. In G. E. Stelmach & V. Homberg (Eds.), *Sensorimotor Impairment in the Elderly* (pp. 73-87). Norwell, MA: Kluwer Academic.

Elkind, J. S., Rubin, E., Rosenthal, S., Skoff, B., & Prather, P. (2001). A simulated reality scenario compared with the computerized Wisconsin card sorting test: an analysis of preliminary results. *Cyberpsychology & Behavior, 4*(4), 489-496.

Farrimond, S., Knight, R. G., & Titov, N. (2006). The effects of aging on remembering intentions: Performance on a simulated shopping task. *Applied Cognitive Psychology, 20,* 533–555.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioural Research Methods, 39*(2), 175-191.

Fazeli, P. L., Ross, L. A., Vance, D. E., & Ball, K. (2013). The relationship between computer experience and computerized cognitive test performance among older adults. *Journal of Gerontology B: Psychological Sciences and Social Sciences, 68(3),* 337-346.

Fozard, J. L., & Wahl, H.-W. (2012). Age and cohort effects in gerontechnology: a reconsideration. *Gerontechnology, 11,* 10–21.

Garden, S., Phillips, L. H., & MacPherson, S. E. (2001). Mid-life aging, open-ended planning, and laboratory measures of executive function. *Neuropsychology, 15(4),* 472-482.

Grady, C. L., & Craik, F. I. M. (2000). Changes in memory processing with age. *Current Opinion in Neurobiology, 10,* 224-231.

Haertel, E.H. (2006). Reliability. In R. L. Brennan, R.L. (Ed.), Educational Measurement, (pp. 64-110). Westport, CT: Greenwood.

Hart, T., Chaparro, B., & Halcomb, C. (2008). Evaluating websites for older adults: adherence to senior-friendly guidelines and end-user performance. *Behaviour & Information Technology, 27(3),* 191–199.

Hawthorn, D. (2007). Interface design and engagement with older people. *Behavior and Information Technology, 26(4),* 333–341.

Hering, A., Cortez, S. A., Kliegel, M., & Altgassen, M. (2014). Revisiting the age-prospective memory-paradox: the role of planning and task experience. *European Journal of Ageing, 11(1),* 99-106.

Jayroe, T. J., & Wolfram, D. (2012). Internet searching, tablet technology and older adults. Proceedings of the American Society for Information Science and Technology, 1-3.

Jovanovski, D., Zakzanis, K., Ruttan, L., Campbell, Z., Erb, S., & Nussbaum, D. (2012). Ecologically Valid Assessment of Executive Dysfunction Using a Novel Virtual Reality Task in Patients with Acquired Brain Injury. *Applied Neuropsychology, 19*(3), 207-220.

Ketcham, C., Seidler, R. D., Van Gemmert, A. W., & Stelmach, G. E. (2002). Age-related kinematic differences as influenced by task difficulty, target size, and movement amplitude. *Journal of Gerontology B, Psychological Sciences and Social Sciences, 57(1),* P54-64.

Kleigel, M., McDaniel, M. A., & Einstein, G. O. (2000). Plan formation, retention, and execution in prospective memory: A new approach and age-related effects. *Memory and Cognition, 28*(6), 1041-1049.

Law, A. S., Trawley, S. L., Brown, L. A., Stephens, A. N., & Logie, R. H. (2012). The impact of working memory load on task execution and online plan adjustment during multitasking in a virtual environment. *The Quarterly Journal of Experimental Psychology, 66(6),* 1241-1258.

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics, 63(8),* 1279-1292.

Levine, B., Stuss, D. T., Milberg, W. P., Alexander, M. P., Schwartz, M., & MacDonald, R. (1998). The effects of focal and diffuse brain damage on strategy application: Evidence from focal lesions, traumatic brain injury and normal aging. *Journal of the International Neuropsychological Society, 4,* 247-264.

Levy, J., & Pashler, H. (2008). Task prioritization in multitasking during driving: Opportunity to abort a concurrent task does not insulate braking responses from dual-task slowing. *Applied Cognitive Psychology, 22*, 507–525.

Logie, R. H., Trawley, S., & Law, A. (2011). Multitasking: multiple, domain-specific cognitive functions in a virtual environment. *Memory & Cognition, 39*, 1561-1574.

MacPherson, S. E., Della Sala, S., & Logie, R. H. (2004). Dual task interference on encoding and retrieval processes in healthy and impaired working memory. *Cortex, 40(1),* 183-184.

MacPherson, S. E., Della Sala, S., Logie, R. H., & Wilcock, G. K. (2007). Specific AD impairment in concurrent performance of two memory tasks. *Cortex, 43*(7), 858-865.

McAlister, C., & Schmitter-Edgecombe, M. (2013). Naturalistic assessment of executive function and everyday multitasking in healthy older adults. *Aging, Neuropsychology, and Cognition, 20(6),* 735-756.

McDowd, J. M., & Craik, F. I. (1988). Effects of aging and task difficulty on divided attention performance. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 267-280.

McGeorge, P., Phillips, L. H., Crawford, J. R., Garden, S. E., Della Sala, S., Milne, A. B., Hamilton, S., & Callender, J. S. (2001). Using virtual environments in the assessment of executive dysfunction. *Presence: Teleoperators and Virtual Environments, 10*(4), 375-383.

Medlock, S., Eslami, S., Askari, M., Sent, D., Buurman, B., De Rooij, S., & Abu-Hanna, A. (2012). Health information seeking behavior of elderly Internet users in the Netherlands. *European Geriatric Medicine, 3,* S90-S91.

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry, 21*(11), 1078-1085.

Mitzner, T. L., Boron, J. B., Fausset, C. B., Adams, A. E., Charness, N., Czaja, S. J., Dijkstra, K., Fisk, A. D., Rogers, W.A., & Sharit, J. (2010). Older adults talk technology: technology usage and attitudes. *Computers in Human Behavior, 26(6),* 1710–1721.

Morey, R. D., & Rouder, J. N. (2015). BayesFactor: 0.9.11-1 CRAN. Zenodo.

Paraskevaides, T., Morgan, C. J. A., Leitza, J. R., Bisby, J. A., Rendell, P. G., & Curran, H. V. (2010). Drinking and future thinking. Acute effects of alcohol on prospective memory and future simulation. *Psychopharmacology, 208,* 301–308.

Pew Internet and American Life Project. (2004). *Older Americans and Internet.* Retrieved April 15, 2005, from http://www.pewinternet.org/pdfs/pip_seniors_online_2004.pdf

Phillips, L. H., Kliegel, M., & Martin, M. (2006). Age and planning tasks: the influence of ecological validity. *International Journal of Aging and Human Development, 62*, 175-184.

Rajendran, G., Law, A. S., Logie, R. H., van der Meulen, M., Fraser, D., & Corley, M. (2011). Investigating multitasking in high-functioning adolescents with autism spectrum disorders using the Virtual Errands Task. *Journal of Autism and Developmental Disorders, 41*, 1445-1454.

Rand, D., Rukan, S. B. A., Weiss, P. L., & Katz, N. (2009). Validation of the Virtual MET as an assessment tool for executive functions. *Neuropsychological Rehabilitation, 19*(4), 583-602.

Rendell, P. G., & Craik, F. I. M. (2000). Virtual week and actual week: Age-related differences in prospective memory. *Applied Cognitive Psychology, 14,* S43–S62.

Rose, N. S., Luo, L., Bialystok, E., Hering, A., Lau, K., & Craik, F. I. (2015). Cognitive processes in the Breakfast Task: Planning and monitoring. *Canadian Journal of Experimental Psychology*.

Rosenthal, R. L. (1967). Covert communication in the psychological experiment. *Psychological Bulletin, 67,* 356–367.

Salthouse, T. A., Rogan, J. D., & Prill, K. A. (1984). Division of attention: Age differences on a visually presented memory task. *Memory and Cognition, 12*, 613-620.

Sanders, C., & Schmitter-Edgecombe, M. (2012). Identifying the nature of impairment in planning ability with normal aging. *Journal of Clinical and Experimental Neuropsychology, 34*(7), 724-737.

Schneider, N., Wilkes, J., Grandt, M., & Schlick, C. M. (2008). Investigation of input devices for the age-differentiated design of human-computer interaction. *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*, 144-148.

Schmitter-Edgecombe, M., McAlister, C., & Weakley, A. (2012). Naturalistic assessment of everyday functioning in individuals with mild cognitive impairment: The Day Out Task. *Neuropsychology, 26, 26(5),* 631-641.

Schnitzspahn, K. M., Ihle, A., Henry, J. D., Rendell, P. G., & Kliegel, M. (2011). The age-prospective memory paradox: An exploration of possible mechanisms. *International Psychogeriatrics, 23(4),* 583–592.

Seidler, R. D., Alberts, J. L., & Stelmach, G. E. (2002). Changes in multi-joint performance with age. *Motor Control, 6*(1), 19-31.

Shallice, T., & Burgess, P. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain, 114,* 727-741.

Smith, M. W., Sharit, J., & Czaja, S. J. (1999). Aging, motor control, and the performance of computer mouse tasks. *Human Factors, 41*(3), 389-396.

Spooner, D. M., & Pachana, N. A. (2006). Ecological validity in neuropsychological assessment: a case for greater consideration in research with neurologically intact populations. *Archives of Clinical Neuropsychology, 21(4),* 327-337.

Titov, N., & Knight, R. G. (2005). A computer-based procedure for assessing functional cognitive skills in patients with neurological injuries: The virtual street. *Brain Injury, 19,* 315-322.

Turner, P., Turner, S., & Van De Walle, G. (2007). How older people account for their experiences with interactive technology. *Behaviour & Information Technology, 26(4),* 287–296.

Wagner, N., Hassanein, K., & Head, M. (2010). Computer use by older adults: A multi-disciplinary review. *Computers in Human Behavior, 26,* 870–882.

Wagner, G. P., & Trentini, C. M. (2009). Assessing executive functions in older adults: A comparison between the manual and the computer-based versions of the Wisconsin Card Sorting Test. *Psychology & Neuroscience, 2(2),* 195–198.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised*. New York: Psychological Corporation.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale-Third Edition*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale-Third Edition*. San Antonio, TX: Psychological Corporation.

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*, 272-292.

Wright, R. E. (1981). Aging, divided attention, and processing capacity. *Journal of Gerontology, 36,* 605-614.

Zhang, L., Abreu, B. C., Seale, G. S., Maasel, B., Christiansen, C. H., & Ottenbacher, K. J. (2003). A virtual reality environment for evaluation of a daily living skill in brain injury rehabilitation: Reliability and validity. *Archives of Physical Medicine and Rehabilitation*, *84*, 1118-1124.

Table 1

Mean (M) scores and standard deviations (SD) for the younger and older adults performing the background measures

	Younger		Older	
	M	*SD*	*M*	*SD*
ACE-R (out of 100)	97.50	2.07	91.75	5.33
Digit span forward (out of 14)	10.20	2.21	7.40	1.81
Digit span backward (out of 14)	9.00	2.10	6.65	2.11
Digit symbol substitution (out of 133)	57.65	7.39	35.95	11.12
Alpha span (out of 14)	6.70	1.53	4.85	1.46

ACE-R = Addenbrooke's Cognitive Examination-Revised

Table 2

Mean (*M*) scores, standard deviations (*SD*), minimum (Min) and maximum scores (Max), skewness and kurtosis for the younger and older adults performing the computerized version of the Breakfast task

	Younger						Older					
	M	*SD*	*Min*	*Max*	*Skewness*	*Kurtosis*	*M*	*SD*	*Min*	*Max*	*Skewness*	*Kurtosis*
Discrepancy[a]	7.20	4.66	2	19	0.34	-0.34	22.25	16.71	4	59	-0.16	-0.73
Range of stop times[a]	19.50	11.45	7	45	0.41	-0.84	72.25	77.61	16	375	0.86	1.66
Deviation from ideal start times[a]	6.80	6.75	1	29	0.48	-0.33	19.20	23.16	5	113	1.21	3.01
Number of checks	19.50	6.26	10	33	0.43	-0.23	13.85	4.98	4	21	-0.15	-0.99
Number of table settings	41.05	10.70	16	60	-0.15	0.15	9.85	5.49	1	20	-0.11	-0.96

[a]ANOVA data log transformed; * Shapiro-Wilks test p < .05

Table 3

Mean (*M*) scores, standard deviations (*SD*), minimum (Min) and maximum scores (Max), skewness and kurtosis for the younger and older adults performing the prop-based version of the Breakfast task

	Younger						Older					
	M	*SD*	*Min*	*Max*	*Skewness*	*Kurtosis*	*M*	*SD*	*Min*	*Max*	*Skewness*	*Kurtosis*
Discrepancy[a]	5.50	5.36	0	19	0.20	0.16	10.20	13.73	2	59	1.06	0.45*
Range of stop times[a]	14.95	22.84	3	106	1.67	2.93*	19.80	19.03	4	65	0.88	0.24*
Deviation from ideal start times[a]	3.60	4.01	0	16	0.35	-0.55	6.60	6.13	0	26	-0.16	-0.14
Number of checks	13.85	6.73	6	38	2.57	8.78*	13.70	4.18	6	20	-0.02	-0.93
Number of table settings	53.00	12.10	24	74	-0.74	0.60	35.20	13.65	18	66	0.54	-0.55

[a]ANOVA data log transformed; * Shapiro-Wilks test p < .05

Table 4

The formulae for calculating the ideal start times for each food.

Food	Ideal start times (in seconds)
Food 1	Start time 1 = 0
Food 4	Start time 2 = 90
Food 2	Start time 3 = (120 + actual start time of food 4 + 30) / 2
Food 5	Start time 4 = (150 + actual start time of food 4 + 60 + actual start time of food 2 + 30) / 3
Food 3	Start time 5 = (210 + actual start time of food 4 +120 + actual start time of food 2 + 90 + actual start time of food 5 + 60) / 4

Table 5

Mean scores (M) and standard deviations (SD) for time spent by the younger and older adults performing the computerized and prop-based versions of the Breakfast task

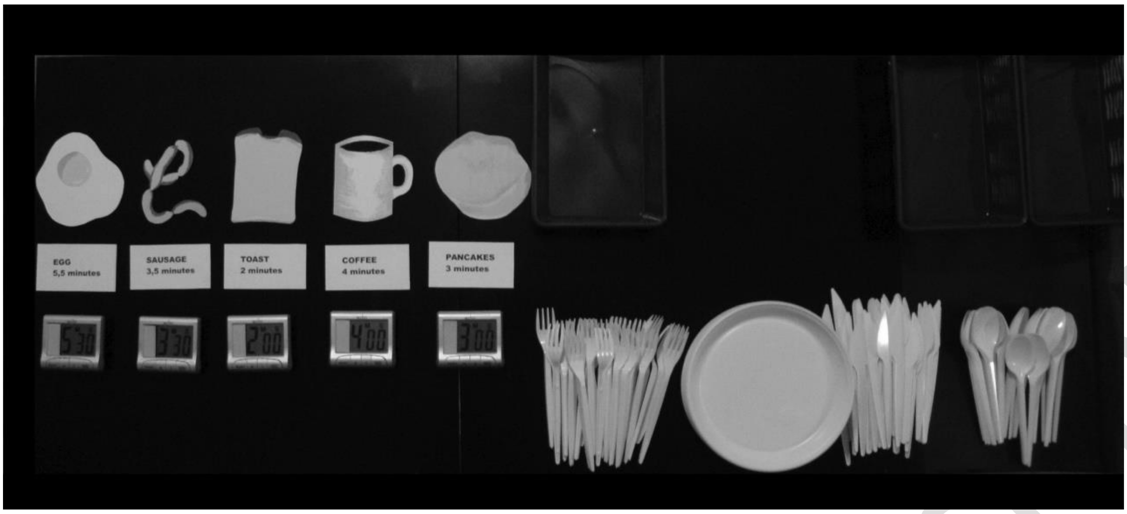
	Younger				Older			
	M	*SD*	*Skewness*	*Kurtosis*	*M*	*SD*	*Skewness*	*Kurtosis*
Computerized								
Total (sec) [a]	361.45	44.00	1.21	2.89	391.63	62.86	0.55	0.86
Setting (sec) [a]	301.70	31.87	0.04	-0.46	302.15	52.82	0.61	1.36
Cooking (sec) [a]	59.75	31.98	0.17	-0.85	89.48	35.16	0.50	-0.45
Prop. Set[a]	0.84	0.06	-0.69	-0.47	0.77	0.07	-0.42	-1.26
Inappropriate	0.10	0.17	0.73	0.83	0.24	0.14	0.98	1.57
Prop-based								
Total (sec) [a]	341.33	10.84	1.79	3.76*	353.70	18.39	1.11	0.28*
Setting (sec) [a]	312.30	13.41	0.49	0.41*	319.15	22.55	0.35	0.00
Cooking (sec) [a]	29.03	6.74	0.21	-0.69	34.55	9.22	0.37	0.75
Prop. Set[a]	0.91	0.02	-0.59	0.16	0.90	0.03	-1.49	3.10*
Inappropriate	0.07	0.05	0.27	-1.32	0.14	0.09	1.14	1.52

Prop. Set. = proportion of total time spent setting table; Inappropriate = proportion of setting time when cooking should have been performed

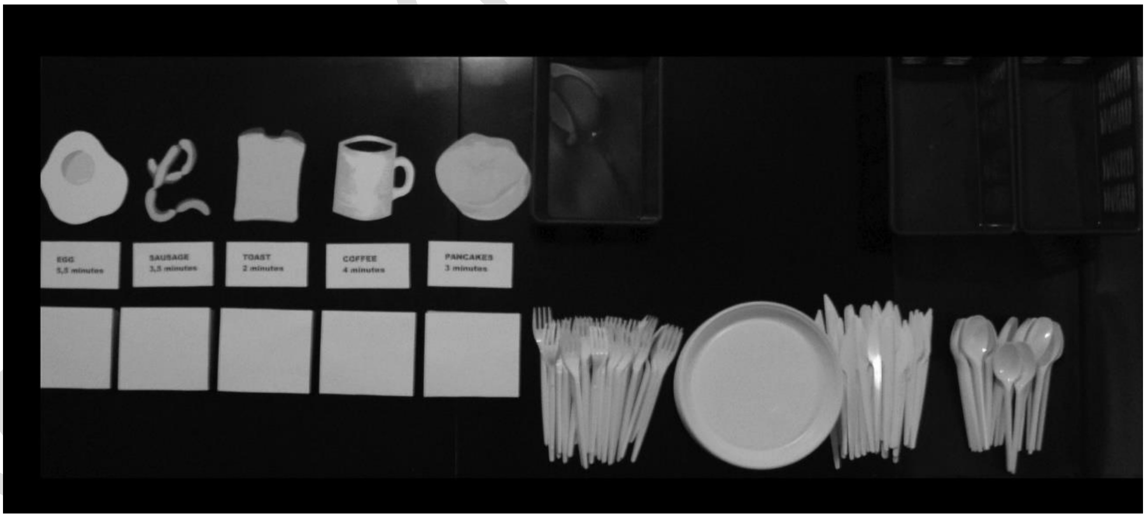
[a]ANOVA data log transformed; * Shapiro-Wilks test p < .05

Figure Captions

Figure 1. Display for the prop-based Breakfast task (top: 1-display condition; middle: 2-display condition; and bottom: 6-display condition). In the 1-display condition, the timers are not covered, in the 2-display condition, the timers are covered but removal of the cover reveals all timers at the same time and in the 6-display condition, the timers are covered and revealed individually. See Figure 1 of Craik & Bialystok (2006) for an example of the computerized Breakfast task.


EGG
5,5 minutes
SAUSAGE
3,5 minutes
TOAST
2 minutes
COFFEE
4 minutes
PANCAKES
3 minutes


EGG
5,5 minutes
SAUSAGE
3,5 minutes
TOAST
2 minutes
COFFEE
4 minutes
PANCAKES
3 minutes


EGG
5,5 minutes
SAUSAGE
3,5 minutes
TOAST
2 minutes
COFFEE
4 minutes
PANCAKES
3 minutes